

# Characters Extraction from Strings on a Document Image Using Handwritten Marks on Touch Screen

Shinnosuke Hattori, Shinji Tsuruoka  
 Graduate School of Regional Innovation Studies  
 Mie University  
 1577 Kurima-Machiya, Tsu, Mie 514-8507, Japan  
 shattori@ip.elec.mie-u.ac.jp

Hiroharu Kawanaka, Haruhiko Takase  
 Graduate School of Engineering  
 Mie University  
 1577 Kurima-Machiya, Tsu, Mie 514-8507, Japan  
 kawanaka@elec.mie-u.ac.jp

**Abstract**—We argued validity of Tablet PCs in the fields of E-learning, and this paper discussed a new character extraction system from strings on the document image using handwritten marks on touch screen. As the first step of this study, we proposed the method to identify handwritten notes/marks to associate strings on the documents with handwritten marks. In this paper, experiments using actual scanned documents were also conducted. As an experimental result, 90% of notes extraction rate was obtained and 60% of marks were appropriately extracted. We confirmed that the proposed method would be a useful user interface for the character extraction from the documents.

**Keywords**-Tablet Device, E-learning, Electronic Documents

## I. INTRODUCTION

Currently, tablet PCs, *e.g.* iPad and Galaxy Tab etc., have been widely used in various situations. For example, we can buy electronic books on-line and read them easily by using tablet PCs. In addition we can also write notes, marks and underlines in the book intuitively. Therefore tablet devices would become indispensable devices.

On the other hand, a lot of researches on E-learning systems are reported [1], [2], [3]. For example, Bonds-Raacke *et al.* conduct a survey of students in the class using tablet PCs, and positive results are obtained [1]. And Ming. Y *et al.* propose a mandarin E-learning system based on speech recognition and evaluation [2]. These are very effective systems for E-learning, however, does not enough capability for actual classes using tablet devices. Actually there are various kinds of contents, *e.g.* sentences, figures, graphs and tables etc., in the textbook and students/lecturers add underlines, marks and notes onto the each page. Of course tablet devices can realize that kind of things easily, because each trajectories of stylus can be obtained as positional data. However, the meanings of inscribed objects heavily depends on their location on the page, thus it is difficult to recognize what the line segment means. For example, if there is a line segment on the sentences, it might mean an underline or mark, but it should be regarded as a note when it is written on the figure or the area without figures/characters. In addition, in the cases of underlines or marks, keywords on

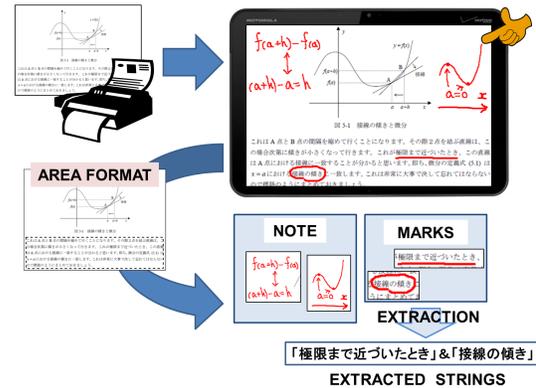


Figure 1. Rough Image of Proposed System

them have to be picked up and recognized to add search tags. Therefore, the method to recognize the meaning of inscribed object is required for E-learning system using tablet devices.

In this study, we discuss techniques to recognize the meaning of inscribed object by analyzing document layout and time-series positional data of trajectories. As the first step of this study, this paper proposes the method to extract strings from inscribed objects on the document and recognize the kind of them. We also show the outline of developed system and describe future works of this study.

## II. METHOD

Figure 1 illustrates the rough image of the proposed system. As the first step, paper-based documents such as textbooks are scanned and converted to electronic documents like PDF. Then we can see the contents of them and add notes/marks on the each page. If users add marks on the page by hand, the system detects it and extracts the strings including keywords corresponding to the marks. For example, if it is an underline, the system extracts strings on the line, and if strings are enclosed a square, the system extract them in the square. On the other hand, an inscribed object on the white area is regarded as a note. In other words, it is important for string extraction to discriminate where users inscribed

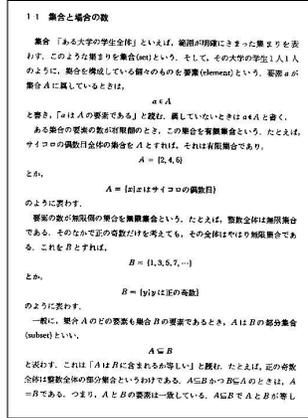


Figure 2. Employed Target Image

an line segment on the page — the area with characters and figures (Printing Area) or other area (Blank Area). Thus the proposed method first applies dilation processing to the scanned image to fill line spaces with black pixels. After this, the regions with black pixels are regarded as printing areas. The number of dilation processing is experimentally determined using the result of preliminary experiments. In this paper, if the inscribed object is in the printing area completely, the object is regarded as the mark and others are did as notes. By using the above criteria, it is expected that an inscribed object on the figure is recognized as a note.

### III. EVALUATION EXPERIMENTS

#### A. Materials

In this study, we conducted evaluation experiments. We used a pen tablet computer (WACOM DTI-520) with Microsoft Windows XP for evaluation. In the experiments, the textbook page shown as Figure 2. are used as the target image. Five examinees were employed and they evaluated the developed system. Each examinees wrote five notes and five marks — the total number of notes and marks was 25, respectively. The subscribed marks was not limited, but most of examinees wrote circle/square marks, underlines on the page.

#### B. Results and Discussion

Table I shows the summary of the obtained experimental results. In the table, success rate, *i.e.* extraction rate, was calculated by the following.

$$\text{Success Rate} = \frac{\# \text{ of Successful Cases}}{\text{Total \# of Notes or Marks}} \times 100 \quad (1)$$

The obtained results indicate that the success rate of mark extraction was lower than that of notes extraction. In the proposed method, segmentation conditions are strict thus the inscribed objects are regarded as notes when they are slightly located on the printing area. As the result, some of marks

Table I  
EXPERIMENTAL RESULT

	Total # of Notes (or Marks)	# of Successful Cases	Success Rate
Notes	25	18	90.0 %
Marks	25	12	60.0 %

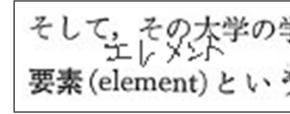


Figure 3. Example Image of Mis-recognize (Rubi)

were regarded as notes even if most of the inscribed objects are in the printing area. In addition, Japanese textbooks have some kinds of characters such as “Hiragana”, “Katakana”, “Chinese Characters” and so on. Some students put rubi (hiragana text) on the Chinese character (or English word) to their easy understanding, and it should be regarded as notes. However, rubi is generally added on the printing area, thus the proposed method mis-recognize the rubi as mark (Figure 3).

### IV. CONCLUSION AND FURTHER WORKS

In this paper, we proposed the method for extracting strings from inscribed mark regions on the document. As a result of experiments, 90% of note extraction rate and 60% of mark extraction rate were obtained. This study is just beginning, thus this algorithm is so simple and immature. The proposed method employ only positional data for extraction now. More investigations and improvements are required for practical use.

### REFERENCES

- [1] Bonds-Raacke, Jennifer M., Raacke, John D., "Using Tablet PCs in the classroom: an investigation of students' expectations and reactions", *Journal of Instructional Psychology*, pp.235 - 239, 2008
- [2] Ming. Y, Bai, ZS, "A Mandarin E-Learning System Based on Speech Recognition and Evaluation", *COMPUTER APPLICATIONS IN ENGINEERING EDUCATION*, pp.651 - 659, 2011
- [3] Guliato D, Boaventura RS, Maia MA, Rangayyan RM, Simedo MS, Macedo TAA, "INDIAM – an e-learning system for the interpretation of mammograms... Interpretation and Diagnosis of Mammograms.", *Journal of Digital Imaging*, pp.405 - 420, 2009
- [4] I. Ota, R. Yamamoto, T. Nishimoto, S. Sagayama, "Online Handwritten Kanji String Recognition Based on Grammar Description of Character Structures", *IEICE PRMU, Pattern Recognition and Media Understanding*, pp.75 - 80, 2008